US009620105B2

(12) **United States Patent**
Mason

(10) **Patent No.:** **US 9,620,105 B2**
(45) **Date of Patent:** **Apr. 11, 2017**

(54) **ANALYZING AUDIO INPUT FOR EFFICIENT SPEECH AND MUSIC RECOGNITION**

(71) Applicant: **Apple Inc.**, Cupertino, CA (US)

(72) Inventor: **Henry Mason**, San Francisco, CA (US)

(73) Assignee: **Apple Inc.**, Cupertino, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 149 days.

(21) Appl. No.: **14/500,740**

(22) Filed: **Sep. 29, 2014**

(65) **Prior Publication Data**

US 2015/0332667 A1 Nov. 19, 2015

Related U.S. Application Data

(60) Provisional application No. 61/993,709, filed on May 15, 2014.

(51) **Int. Cl.**
G10L 25/81 (2013.01)
G10L 15/02 (2006.01)
G10L 25/03 (2013.01)

(52) **U.S. Cl.**
CPC *G10L 15/02* (2013.01); *G10L 25/03* (2013.01); *G10L 25/81* (2013.01); *G10L 2015/025* (2013.01)

(58) **Field of Classification Search**
CPC G10L 15/02; G10L 25/81
USPC 704/249
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

1,559,320 A	10/1925	Hirsh	
2,180,522 A	11/1939	Henne	
3,704,345 A	11/1972	Coker et al.	
3,710,321 A	1/1973	Rubenstein	
3,828,132 A	8/1974	Flanagan et al.	
3,979,557 A	9/1976	Schulman et al.	
4,013,085 A	3/1977	Wright	
4,081,631 A	3/1978	Feder	

(Continued)

FOREIGN PATENT DOCUMENTS

CH	681573 A5	4/1993
CN	1673939 A	9/2005

(Continued)

OTHER PUBLICATIONS

"Top 10 Best Practices for Voice User Interface Design" available at <http://www.developer.com/voice/article.php/1567051/Top-10-Best-Practices-for-Voice-UserInterface-Design.htm>, Nov. 1, 2002, 4 pages.

(Continued)

Primary Examiner — Susan McFadden
(74) *Attorney, Agent, or Firm* — Morrison & Foerster LLP

(57) **ABSTRACT**

Systems and processes for analyzing audio input for efficient speech and music recognition are provided. In one example process, an audio input can be received. A determination can be made as to whether the audio input includes music. In addition, a determination can be made as to whether the audio input includes speech. In response to determining that the audio input includes music, an acoustic fingerprint representing a portion of the audio input that includes music is generated. In response to determining that the audio input includes speech rather than music, an end-point of a speech utterance of the audio input is identified.

60 Claims, 4 Drawing Sheets